



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

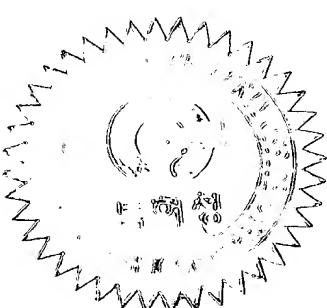
This is to certify that the following application annexed hereto  
is a true copy from the records of the Korean Industrial  
Property Office.

출 원 번 호 : 특허출원 1999년 제 43712 호  
Application Number

출 원 년 월 일 : 1999년 10월 11일  
Date of Application

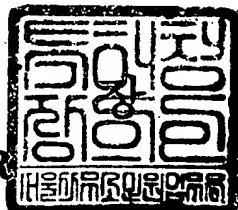
출 원 인 : 한국전자통신연구원  
Applicant(s)

2000 년 09 월 20 일



특 허 청

COMMISSIONER



CERTIFIED COPY OF  
PRIORITY DOCUMENT

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【제출일자】	1999. 10. 11
【발명의 명칭】	효율적인 비디오 개관 및 브라우징을 위한 요약 비디오 기술구조와 그 기술구조를 이용한 요약 비디오 기술 데이터 생성 방법
【발명의 영문명칭】	A Skim Video Description Scheme and Description Browsing Method for Efficient Overview and Browsing
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【성명】	전영일
【대리인코드】	9-1998-000540-4
【포괄위임등록번호】	1999-054594-1
【발명자】	
【성명의 국문표기】	김재곤
【성명의 영문표기】	KIM, Jae Gon
【주민등록번호】	670726-1788014
【우편번호】	302-120
【주소】	대전광역시 서구 둔산동 908
【국적】	KR
【발명자】	
【성명의 국문표기】	장현성
【성명의 영문표기】	CHANG, Hyun Sung
【주민등록번호】	741203-1024526
【우편번호】	302-150
【주소】	대전광역시 서구 만년동 강변아파트 112-506
【국적】	KR
【발명자】	
【성명의 국문표기】	김문철
【성명의 영문표기】	KIM, Mun Chul
【주민등록번호】	670104-1691416

【우편번호】	302-122		
【주소】	대전광역시 서구 둔산2동 912번지 등지아파트 110동 405호		
【국적】	KR		
【발명자】			
【성명의 국문표기】	김진웅		
【성명의 영문표기】	KIM, Jin Woong		
【주민등록번호】	591223-1011621		
【우편번호】	305-390		
【주소】	대전광역시 유성구 전민동 엑스포아파트 305동 1603호		
【국적】	KR		
【신규성주장】			
【공개형태】	간행물 발표		
【공개일자】	1999.10.01		
【심사청구】	청구		
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사 를 청구합니다. 대리인 전영일 (인)		
【수수료】			
【기본출원료】	20	면	29,000 원
【가산출원료】	11	면	11,000 원
【우선권주장료】	0	건	0 원
【심사청구료】	24	항	877,000 원
【합계】	917,000 원		
【감면 사유】	정부출연연구기관		
【감면후 수수료】	458,500 원		
【첨부서류】	1. 요약서·명세서(도면)_1통 2. 신규성(출원시의 특례)규정 을 적용받기 위한 증명서류_1통		

### 【요약서】

#### 【요약】

본 발명은 주문형 요약 비디오 등에 적합한 요약 비디오 기술 구조와 그 생성 방법에 관한 것이다. 본 발명에 따른, 요약 비디오를 기술하기 위한 계층적인 구조를 갖는 요약 기술 구조(Summary Description Scheme)는, 하이라이트 레벨(Highlight Level)에 대해 기술하는 하나 이상의 다수의 레벨을 가질 수 있는 하이라이트 레벨 기술 구조(Highlight Level Description Scheme)와 상기 하이라이트 레벨(Highlight Level)의 의미와 상기 하이라이트 레벨들 간의 관계에 대한 정보를 기술하는 계층 유형(Hierarchy Type)을 포함하는 계층적 요약 기술 구조(Hierarchical Summary Description Scheme)로 이루어지는데, 상기 하이라이트 레벨 기술 구조(Highlight Level Description Scheme)는 요약 비디오 구간에 해당하는 정보를 기술하는 하나 이상의 하이라이트 세그먼트 기술 구조(Highlight Segment Description Scheme)를 포함하고, 상기 하이라이트 세그먼트 기술 구조(Highlight Segment Description Scheme)는 요약 비디오 구간의 시간 정보를 기술하거나 요약 비디오 구간의 비디오 데이터 정보를 기술하는 세그먼트 위치지정 기술 구조(Segment Locator Description Scheme)를 포함하여 이루어지는 것을 특징으로 하는 요약 기술 구조를 포함한다.

#### 【대표도】

도 2

### 【명세서】

#### 【발명의 명칭】

효율적인 비디오 개관 및 브라우징을 위한 요약 비디오 기술구조와 그 기술구조를 이용한 요약 비디오 기술 데이터 생성 방법 { A Skim Video Description Scheme and Description Browsing Method for Efficient Overview and Browsing }

#### 【도면의 간단한 설명】

도 1은 본 발명에 따른 기술구조(description scheme)에 따라서 요약 비디오 기술데이터를 생성하기 위한 흐름도이고,

도 2는 본 발명에 따른 요약 비디오 기술데이터를 기술하기 위한 계층적 기술 구조의 자료구조를 UML(Unified Modeling Language)로 도시한 것이고,

도 3은 본 발명에 따른 요약 비디오 재현 및 브라우징 툴의 사용자 인터페이스의 일 실시예이고,

도 4는 본 발명에 따른 요약 비디오 기술 데이터를 이용한 계층적 브라우징을 위한 데이터 및 제어 흐름에 대한 구성도이다.

#### 【발명의 상세한 설명】

##### 【발명의 목적】

##### 【발명이 속하는 기술분야 및 그 분야의 종래기술】

<5> 본 발명은 주문형 요약 비디오 등에 적합한 요약 비디오 기술 구조와 그 생성 방법에 관한 것으로서, 특히 요약 비디오와 함께 요약 비디오에 포함된 각 비디오 구간마다 대

표 프레임 정보, 분류 정보, 대표 음향 정보를 포함하는 비디오 기술구조와 그 생성 방법에 관한 것이다.

- <6> 본 발명이 속하는 기술분야는 내용기반 비디오 데이터 인덱싱 및 브라우징/검색 분야로 비디오를 내용기반(content based)으로 요약하고 이를 기술하는 분야이다. 비디오를 요약하는 방법은 크게 동적 요약(dynamic summary)과 정적 요약(static summary)으로 나눌 수 있는데 본 발명에 따른 비디오 기술 구조는 동적 요약을 위한 것이다.
- <7> 일반적으로, 기존의 요약 비디오 및 기술구조는 단순히 요약 비디오에 포함된 비디오 구간에 대한 정보만을 제공함으로써 요약 비디오의 재현을 통하여 전체 비디오의 내용을 전달하는데 국한된다. 그러나 많은 경우 요약 비디오를 통해서 전체 내용을 개관하는데 그치기보다는 전체 내용의 개관을 통하여 관심 있는 부분을 다시 확인하기 위한 브라우징이 필요하다.
- <8> 또한 기존의 요약 비디오는 요약 비디오 제공자가 정한 기준에 의해서 중요하다고 판단되는 비디오 구간만을 사용자에게 제공한다. 따라서 사용자와 비디오 제공자의 기준이 다른 경우, 혹은 사용자가 특별히 원하는 기준이 있을 경우 사용자는 원하는 형태의 요약 비디오를 얻을 수 없다. 즉, 기존의 요약 비디오는 몇 가지 레벨의 요약 비디오가 제공되어 사용자가 원하는 레벨의 요약 비디오를 선택하도록 하지만 요약 비디오의 내용에 따른 선택을 할 수 없으므로 사용자의 선택 범위가 제한적이다.
- <9> 발명의 명칭이 'method and apparatus for video browsing based on content'

and structure' 이고 등록 번호가 US5821945 인 특허에서는 비디오를 간략히 표현하고 그 표현을 통하여 원하는 내용의 비디오로 접근하는 브라우징 기능을 제공한다. 그러나 대표 프레임에 기반한 정적인 요약이고 기존의 정적 요약은 비디오 샷(shot)의 대표 프레임을 이용하여 요약하는데, 대표 프레임은 단지 그 샷을 대표하는 영상 정보만 제공하므로 요약을 이용한 정보 전달에 한계가 있다. 이에 의해 본 발명에 따른 비디오 기술구조와 브라우징 방법은 비디오 세그먼트에 기반한 동적 요약을 이용한다.

<10> 1999년에 ISO/IEC JTC1/SC29/WG11 MPEG-7 Output Document No. N2844 에 발표된 'MPEG-7 Description Scheme (V0.5)'에서 제안된 비디오 기술구조는 동적 요약 비디오의 각 비디오 세그먼트의 구간 정보만을 기술한다. 이에 의해 본 발명에 따른 비디오 기술구조는 동적 요약 비디오와 함께 요약 비디오의 재현을 바탕으로 브라우징과 주문형 요약 비디오를 제공하는 기능을 가지고 있다. 또한, 본 발명에 따른 비디오 기술 구조는 대표 프레임에 그 대표 프레임에 해당하는 비디오 구간의 대표 음향을 함께 기술함으로써 더 많은 정보로써 요약을 할 수 있도록 한다.

### 【발명이 이루고자 하는 기술적 과제】

<11> 따라서, 본 발명은 상기의 문제점을 개선하기 위하여 요약 비디오와 함께 요약 비디오에 포함된 각 비디오 구간마다 대표 프레임 정보, 분류 정보, 대표 음향 정보를 포함하고, 요약 비디오의 내용에 대한 사용자의 선택을 제공하는 주문형 요약과 효과적인 브라우징을 가능하게 하는 계층적 비디오 기술구조와 그 기술구조를 이용한 요약 비디오 기술 데이터 생성 방법을 제공하는 것을 목표로 한다. 또한, 요약 비디오에 포함된 각 비디오 구간의 대표 프레임과 함께 대표 음향을 포함함으로써 더 효과적인 내용 파악을 통

한 브라우징 기능을 제공한다

### 【발명의 구성 및 작용】

- <12> 이와 같은 목적을 달성하기 위하여 본 발명은, 요약비디오를 기술하는 계층적 요약 기술구조에 있어서, 적어도 하나 이상의 요약 비디오구간들의 요약비디오구간정보를 각각 기술하는 상기 요약 비디오구간 수만큼의 하이라이트 세그먼트기술구조를 포함하는 하이라이트레벨 기술구조와, 상기 하이라이트레벨의 의미를 기술하는 계층유형을 포함하고, 상기 하이라이트 세그먼트기술구조는 요약 비디오구간의 시간정보를 기술하거나 요약 비디오구간의 비디오 데이터정보를 기술하는 세그먼트 위치지정기술구조를 포함하여 이루어진 것을 특징으로 한다.
- <13> 바람직하게는, 상기 하이라이트레벨 기술구조는 다수의 하이라이트레벨을 포함하고, 상기 계층유형은 상기 하이라이트레벨의 레벨간의 관계를 더 포함한 것을 특징으로 한다.
- <14> 바람직하게는, 상기 계층적 요약기술구조는, 요약비디오구간을 선택하는 데 기반이 된 사건들의 종류를 기술하는 요약기준기술구조를 더 포함한 것을 특징으로 한다.
- <15> 바람직하게는, 상기 하이라이트 세그먼트 기술구조는, 비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조를 더 포함한 것을 특징으로 한다.
- <16> 바람직하게는, 상기 영상위치지정기술구조는, 대표프레임에 대한 시간구간정보를 기술하는 것을 특징으로 한다.
- <17> 바람직하게는, 상기 영상위치지정기술구조는, 대표프레임의 영상데이터를 기술하

는 것을 특징으로 한다.

- <18> 바람직 하게는, 상기 하이라이트 세그먼트기술구조는, 비디오구간을 대표하는 음향 정보를 기술하는 대표음향기술구조를 더 포함한 것을 특징으로 한다.
- <19> 바람직 하게는, 상기 대표음향기술구조는, 대표음향에 대한 음향시간구간정보를 기술하는 것을 특징으로 한다.
- <20> 바람직 하게는, 상기 대표음향기술구조는, 대표음향의 음향데이터를 기술하는 것을 특징으로 한다.
- <21> 바람직 하게는, 상기 하이라이트세그먼트기술구조는, 요약비디오구간의 사건종류를 기술하는 세그먼트유형을 더 포함한 것을 특징으로 한다.
- <22> 바람직 하게는, 상기 하이라이트세그먼트기술구조는, 비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조, 비디오구간을 대표하는 음향 정보를 기술하는 대표음향기술구조, 및 요약비디오구간의 사건종류를 기술하는 세그먼트유형을 더 포함한 것을 특징으로 한다.
- <23> 본 발명은, 적어도 하나 이상의 요약 비디오구간들의 요약비디오구간정보를 각각 기술하는 상기 요약 비디오구간 수만큼의 하이라이트 세그먼트기술구조를 포함하는 하이라이트레벨 기술구조와, 상기 하이라이트레벨의 의미를 기술하는 계층유형을 포함하고, 상기 하이라이트 세그먼트기술구조는, 요약 비디오구간의 시간정보를 기술하거나 요약 비디오구간의 비디오 데이터정보를 기술하는 세그먼트 위치지정기술구조와, 상기 요약비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조와, 상기 요약비디오구간의 사

건종류를 기술하는 세그먼트유형구조를 포함하는 데이터가 기록된 컴퓨터로 읽을 수 있는 기록매체를 포함한다.

- <24> 본 발명은, 원 비디오를 입력받아 요약 기술구조에 따라서 요약 비디오기술 데이터를 생성하는 요약 비디오 기술데이터 생성방법에 있어서, 원 비디오를 입력받고 분석하여 비디오 분석결과를 출력하는 비디오 분석단계와; 요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의단계; 상기 원 비디오 분석 결과와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내용을 요약할 수 있는 비디오 구간을 선택하여 요약 비디오구간정보를 구성하는 요약비디오 구간선택단계; 및 상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술단계를 포함하여 이루어진 것을 특징으로 한다.
- <25> 바람직 하게는, 상기 비디오 분석단계는, 원 비디오를 입력으로 받아들여 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출단계와, 상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출단계, 상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출단계를 포함하여 이루어진 것을 특징으로 한다.
- <26> 바람직 하게는, 상기 요약규칙정의단계는 요약비디오구간을 선택하는데 기반이 되는 요약사건종류를 정의하여 상기 요약비디오기술단계에 제공하는 것을 특징으로 한다.
- <27> 바람직 하게는, 상기 요약 비디오 구간정보를 입력받아 대표 프레임을 추출하여 상기 요약비디오기술단계에 제공하는 대표프레임 추출단계를 더 포함한 것을 특징으로 한다.

- <28> 바람직 하게는, 상기 요약 비디오 구간정보를 입력받아 대표음향을 추출하여 상기 요약비디오기술단계에 제공하는 대표음향추출단계를 더 포함한 것을 특징으로 한다.
- <29> 본 발명은, 원 비디오를 입력받아 요약 기술구조에 따라서 요약 비디오기술 데이터를 생성하는 요약 비디오 기술데이터 생성시스템에 있어서, 원 비디오를 입력받고 분석하여 비디오 분석결과를 출력하는 비디오 분석수단과; 요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의수단; 상기 원 비디오 분석 결과와 상기 요약 규칙을 입력받아 원 비디오로부터 비디오내용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택수단; 및 상기 요약비디오 구간선택수단에서 정의된 요약비디오구간정보를 입력받아 계층적요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술수단을 포함하여 이루어진 것을 특징으로 한다.
- <30> 바람직 하게는, 상기 비디오 분석수단은, 원 비디오를 입력으로 받아들여 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출수단과, 상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출수단, 상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출수단을 포함하여 이루어진 것을 특징으로 한다.
- <31> 바람직 하게는, 상기 요약규칙정의수단은 요약비디오구간을 선택하는데 기반이 되는 요약사건종류를 정의하여 상기 요약비디오기술수단에 제공하는 것을 특징으로 한다.
- <32> 바람직 하게는, 상기 요약비디오구간선택수단의 요약 비디오 구간정보를 입력받아 대표 프레임을 추출하여 상기 요약비디오기술수단에 제공하는 대표프레임 추출수단을 더 포함한 것을 특징으로 한다.

<33> 바람직하게는, 상기 요약비디오구간선택수단의 요약 비디오 구간정보를 입력받아 대표음향을 추출하여 상기 요약비디오기술수단에 제공하는 대표음향추출수단을 더 포함한 것을 특징으로 한다.

<34> 본 발명은, 컴퓨터에, 입력되는 원 비디오에서 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출단계와, 상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출단계, 상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출단계 요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의단계; 상기 검출된 에피소드와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택단계; 및 상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술단계를 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 포함한다.

<35> 본 발명은, 컴퓨터에, 입력되는 원 비디오에서 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출수단과, 상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출수단, 상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출수단, 요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의수단, 상기 검출된 에피소드와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를

구성하는 요약비디오 구간선택수단, 및 상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술수단을 기능시키기 위한 컴퓨터로 읽을 수 있는 기록매체를 포함한다.

- <36> 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 일 실시예를 상세하게 설명한다.
- <37> 도 1은 본 발명에 따른 기술구조(description scheme)에 따라서 요약 비디오 기술 데이터를 생성하기 위한 흐름도이다. 도 1에 도시된 바와 같이, 본 발명에 따른 요약 비디오 기술 데이터 생성 방법은 특징 추출 단계(101), 사건 검출 단계(102), 에피소드 검출 단계(103), 요약 비디오 구간 선택 단계(104), 요약 규칙 정의 단계(105), 대표 프레임 추출 단계(106), 대표 음향 추출 단계(107), 요약 비디오 기술 단계(108)로 구성된다.
- <38> 특징 추출 단계(101)는 원 비디오를 입력하여 요약 비디오를 생성하기 위하여 필요한 특징들을 추출한다. 일반적인 특징으로는 샷 경계, 카메라 움직임, 자막 영역, 얼굴 영역 등이 있다. 특징 추출 단계에서는 이를 특징들을 추출하여 특징의 종류와 이를 특징들이 검출되는 비디오 시간 구간을 (특징 종류, 특징 일련번호, 시간구간)의 형태로 사건 검출 단계로 출력한다. 예를 들면 카메라 움직임의 경우 (카메라줌 1, 100~150)은 카메라 줌 1번(첫 번째 줌)이 100~150 프레임에서 검출되었다는 정보를 표현한다.
- <39> 사건 검출단계(102)에서는 원 비디오에 포함된 주요 사건들을 검출한다. 이를 사건은 원 비디오의 내용을 잘 표현할 수 있어야 하고 요약 비디오를 생성하는데 기준이 되는 것들이기 때문에 일반적으로 원 비디오의 양쪽에 따라서 다르게 정의된다. 사건은 상위의 의미 레벨을 나타낼 수도 있고 상위의 의미를 직접 유추할 수 있는 비주얼 특징일

수도 있다. 예를 들어, 축구 비디오의 경우 골, 슛, 자막, 재재생(replay) 등을 사건으로 정의할 수 있다.

<40> 사건 검출 단계(102)는 검출한 사건의 종류와 그 시간 구간을 (사건 종류, 사건번호, 시간 구간) 형태로 출력한다. 예를 들면 (골 1, 200~300) 형태로 첫 번째 골이 200~300 프레임 사이에 발생했다는 사건 정보를 출력한다.

<41> 에피소드 검출 단계(103)는 검출된 사건을 바탕으로 비디오를 이야기 흐름을 기반으로 구분한다. 주요 사건을 검출한 다음 그 주요 사건을 중심으로 그 사건에 따른 부대 사건을 포함하여 하나의 에피소드로 검출한다. 일례로 축구 비디오의 경우 골과 슛은 주요 사건이 되고 그 사건의 부대 사건으로 골이나 슛이 발생했을 때의 벤치 장면, 관중석 장면, 골 세레모니 장면, 골 장면의 재재생 등이 그 사건의 부대 사건을 구성한다. 즉 골과 슛을 중심으로 에피소드를 검출한다.

<42> 에피소드 검출 정보는 (에피소드 번호, 시간 구간, 우선 순위, 특징 샷, 사건 연결 정보)의 형태로 출력한다. 여기서 에피소드 번호는 에피소드의 일련 번호이고 시간 구간은 그 에피소드의 시간 구간을 샷 단위로 나타낸다. 우선 순위는 그 에피소드의 중요도를 나타낸다. 특징 샷은 그 에피소드를 구성하는 샷들 중에서 가장 중요한 정보를 포함한 샷 번호를 나타내고 사건 연결 정보는 그 에피소드와 관련된 사건들의 사건 번호를 나타낸다. 예를 들면 (에피소드1, 4~6, 1, 5, 골 1, 자막3)으로 표시할 경우, 첫 번째 에피소드는 4~6번째 샷을 포함하고, 우선 순위는 높고(1), 특징 샷은 5번 샷이고 연결된 이벤트는 1번 골과 3번 자막임을 나타낸다.

<43> 요약 비디오 구간 선택 단계(104)는 검출된 에피소드를 바탕으로 원 비디오 내용을 잘 요약할 수 있는 비디오 구간을 선택한다. 이 구간 선택 기준은 미리 정해진 요약 규

칙 정의 단계(105)의 요약 규칙에 따라서 수행한다.

- <44>      요약 규칙 정의 단계(105)에서는 요약 구간을 선택하기 위한 규칙을 정의하고 요약 구간을 선택하기 위한 제어신호를 출력한다. 또한 요약 규칙 정의 단계(105)에서는 요약 비디오 구간을 선택하는데 기반이 되는 요약 사건 종류를 요약 비디오 기술 단계(108)로 출력한다.
- <45>      요약 비디오 구간 선택 단계는 선택된 요약 비디오의 구간들의 시간 정보와 요약 비디오 구간의 구간 사건 종류를 프레임 단위로 출력한다. 즉 (100~200, 골), (500~700, 슛) ..... 의 형태로 요약 비디오의 구간으로 선택된 비디오 세그먼트는 100~200 프레임, 500~700 프레임 ... 이고 각 세그먼트의 사건은 골과 슛임을 나타낸다.
- <46>      요약 비디오 구간 선택이 완료되면 그 요약 비디오 구간 정보를 이용하여 그 비디오 구간에서의 대표 프레임과 대표 음향을 대표 프레임 추출 단계(106)와 대표 음향 추출 단계(107)에서 각각 추출한다.
- <47>      대표 프레임 추출단계는 그 요약 비디오 구간을 대표하는 영상의 프레임 번호 또는 그 영상 데이터를 출력한다. 대표 음향부는 그 요약 비디오 구간을 대표하는 음향 데이터 또는 음향 시간 구간을 출력한다.
- <48>      대표 프레임은 그 비디오 구간을 대표하는 한 프레임의 영상 정보이고 대표 음향은 그 구간의 대표하는 음향이다. 일례로 축구 비디오의 경우 선수 이름이나 슛, 골, 센터 링 등의 축구 용어 또는 관중들의 함성 등이 대표 음향이 될 수 있다.
- <49>      요약 비디오 기술 단계(108)에서는 도 2에 기술된 본 발명에 따른 계층적 기술구조(Hierarchical Summary Description Scheme)에 따라서 효과적인 요약과 브라우

정 기능이 가능하도록 관련 정보를 기술한다. 여기에 포함되는 주요 정보는 요약 비디오의 요약 사건 종류와 각 요약 비디오 구간을 기술하는 정보로 시간 정보, 대표프레임, 대표 음향, 구간 사건 종류 정보들이다.

<50>      요약 비디오 기술단계는 도 2의 기술 구조에 따른 요약 비디오 기술 데이터를 출력 한다.

<51>      도 2는 본 발명에 따른 요약 비디오 기술 데이터를 기술하기 위한 계층적 요약 기술 구조(Hierarchical summary Description Scheme)의 자료구조를 UML(Unified Modeling Language)로 도시한 것이다.

<52>      요약 비디오를 기술하는 계층적 요약 기술 구조(Hierarchical summary Description Scheme) (201)는 하나의 하이라이트 레벨 기술 구조(Highlight level Description Scheme) (202), 하나의 계층 유형(Hierarchy type) (203), 하나 또는 제로 개의 요약 기준 기술 구조(Summary criterion Description Scheme) (204)로 구성된다. 여기서 하이라이트 레벨 기술 구조(Highlight level Description Scheme)는 다수의 레벨을 가질 수 있다. 계층 유형(Hierarchy type)은 하이라이트 레벨의 의미와 레벨간의 관계에 대한 정보를 기술한다. 요약 기준 기술 구조(Summary criterion Description Scheme)는 기술하는 요약 비디오의 요약 비디오 구간 종류에 대한 정보를 기술하는 것으로 요약 비디오 구간을 선택하는데 기반이 된 사건의 종류들을 나열하여 기술한다.

<53>      하이라이트 레벨 기술 구조(Highlight level Description Scheme) (202)는 그 레벨의 요약 비디오를 구성하는 비디오 구간 수만큼의 하이라이트 세그먼트 기술 구조(Highlight segment Description Scheme) (205)로 구성된다. 하이라이트 세그먼트 기술 구조(Highlight segment Description Scheme)는 각 요약 비디오 구간에 해당하는 정보를

기술한다. 하이라이트 세그먼트 기술 구조(Highlight segment Description Scheme)는 한 개의 세그먼트 위치지정 기술 구조(Segment locator Description Scheme) (206)와 제로 또는 다 수의 영상 위치지정 기술 구조(Image locator Description Scheme) (207), 그리고 제로 또는 한 개의 대표 음향 기술 구조(Key sound Description Scheme) (208), 세그먼트 유형(Segment type) (209)으로 구성된다. 세그먼트 위치지정 기술 구조(Segment locator Description Scheme)는 해당 비디오 구간의 시간 정보를 기술하거나 해당 비디오 구간의 비디오 데이터 정보를 기술하여 원 비디오에서 해당 구간으로 접근할 수 있게 한다. 영상 위치지정 기술 구조(Image locator Description Scheme)는 그 해당 비디오 구간을 대표하는 대표 프레임에 관한 시간 정보 등을 기술하거나 대표 프레임의 영상 데이터를 기술하여 요약 비디오를 바탕으로 효과적인 브라우징 기능을 제공하도록 한다. 대표 음향 기술 구조(Key sound Description Scheme)는 그 해당 비디오 구간을 대표하는 음향 정보를 음향 데이터 또는 음향 시간 구간으로 기술하여 대표 프레임과 함께 그 해당 비디오 구간을 효과적으로 표현하도록 하여 브라우징 기능을 개선한다. 세그먼트 유형(Segment type)은 그 해당 요약 비디오 구간에 해당하는 사건의 종류를 나타내는 구간 사건 종류를 기술하여 사용자 주문형 요약 비디오를 제공하도록 한다.

<54>        도 3은 요약 비디오 재현 및 브라우징 틀의 사용자 인터페이스의 구성도이다. 비디오 재현부(301)는 사용자의 제어에 따라서 원 비디오나 또는 요약 비디오를 재현한다. 원 비디오 대표 프레임부(305)는 원 비디오의 샷들의 대표 프레임 재현한다. 즉 일련의 축소된 크기의 영상들로 구성된다. 사용자는 대표 프레임을 클

릭하여 대표 프레임에 해당하는 원 비디오의 샷으로 접근한다. 요약 비디오 레벨0 대표 프레임부 및 대표 음향부(307)와 요약 비디오 레벨1 대표 프레임부 및 대표 음향부(306)는 각각 요약 비디오 레벨0과 요약 비디오 레벨1의 각 비디오 구간을 대표하는 프레임과 음향 정보를 재현한다. 즉 일련의 축소된 크기의 영상과 음향을 나타내는 마크 영상으로 구성된다. 사용자가 요약 비디오 대표 프레임부 및 대표 음향부의 대표 프레임을 클릭하면 그 대표 프레임에 해당하는 원 비디오 구간으로 접근한다. 이 때 요약 비디오의 대표 프레임에 해당하는 대표 음향 마크를 클릭하면 그 비디오 구간의 대표 음향이 재현된다.

<55> 주문형 요약 비디오 제어부(302)는 주문형 요약 비디오 생성을 위하여 사용자의 선택을 위한 제어를 입력한다. 사용자는 레벨선택부(303)를 통하여 요약 비디오의 레벨을 선택하고 내용선택부(304)를 통하여 제공되는 요약 비디오 구간의 종류 중에서 원하는 종류의 요약 비디오를 생성하여 재현한다.

<56> 도 4는 본 발명의 요약 비디오를 이용한 계층적 브라우징을 위한 데이터 및 제어 흐름에 대한 구성도이다. 브라우징은 도 3의 사용자 인터페이스를 이용하여 브라우징을 위한 데이터들을 도 4의 방법으로 접근하여 수행한다. 브라우징을 위한 데이터들은 요약 비디오와 요약 비디오의 대표 프레임, 원 비디오(406)와 원 비디오 대표 프레임(405)이다. 요약 비디오는 두 개의 레벨을 갖는 것으로 한다. 물론 두 개 이상의 레벨을 가질 수도 있다. 요약 비디오 레벨0(401)은 요약 비디오 레벨1(403) 보다 더 짧게 요약된 것이다. 즉, 요약 비디오 레벨1이 요약 비디오 레벨0 보다 더 많은 내용을 포함하고 있다. 요약 비디오 레벨0 대표 프레임(402)은 요약 비디오 레벨0의 대표 프레임이고, 요약 비디오 레벨1 대표 프레임(404)은 요약 비디오 레벨1의 대표 프레임이다.

<57>      요약 비디오와 원 비디오는 도 3의 비디오 재현부(301)를 통하여 재현된다. 요약 비디오 레벨0 대표 프레임은 요약 비디오 레벨0 대표 프레임부 및 대표 음향부(306)에 표시되고, 요약 비디오 레벨1 대표 프레임은 요약 비디오 레벨1 대표 프레임부 및 대표 음향부(307)에 표시된다. 원 비디오 대표 프레임은 원 비디오 대표 프레임부(305)에 표시된다.

<58>      도 4에 도시된 본 발명의 계층적 브라우징 방법은 다음과 같은 몇 가지의 계층적 경로를 갖는다.

<59>      경우 1) (1) - (3)

<60>      경우 2) (1) - (4) - (5)

<61>      경우 3) (1) - (4) - (6) - (7)

<62>      경우 4) (2) - (5)

<63>      경우 5) (2) - (6) - (7)

<64>      전체적인 브라우징 기법은 다음과 같다. 먼저 원 비디오의 요약 비디오를 재현해서 원 비디오의 전체 내용을 파악한다. 이 때 요약 비디오는 요약 비디오 레벨0을 재현할 수도 있고 요약 비디오 레벨1을 재현할 수도 있다. 요약 비디오를 재현한 다음 요약 비디오에서 더 자세히 브라우징 하고자 할 때 관심 있는 비디오 구간을 요약 비디오 대표 프레임을 통하여 확인한다. 정확히 찾고자 하는 장면이 요약 비디오 대표 프레임에서 확인이 되면 그 대표 프레임을 연결된 원 비디오의 비디오 구간으로 바로 접근하여 재생한다. 그렇지 않은 경우 좀더 자세한 정보가 필요한 경우 다음 레벨의 대표 프레임을 파악하거나 원 비디오의 대표 프레임의 내용을 계층적으로 파악하여 원하는 원 비디오로 접근

근한다. 이러한 계층적 브라우징 기법은 원하는 내용을 접근하기 위하여 원 비디오를 재생하면서 브라우징 하면 많은 시간이 걸릴 수 있는데 원 비디오의 내용을 계층화된 대표 프레임을 통해서 바로 접근하므로 브라우징 시간을 최대한 줄일 수 있다.

<65> 경우 1)은 요약 비디오 레벨0을 재현하고 요약 비디오 레벨0 대표 프레임으로부터 바로 원 비디오로 접근하는 경우이다. 경우 2)는 요약 비디오 레벨0을 재현하고 요약 비디오 레벨0 대표 프레임에서 가장 관심 있는 대표 프레임을 선택하고 원 비디오에 접근하기 전에 더 자세한 정보를 파악하기 위해서 그 대표 프레임 근처에 해당하는 요약 비디오 레벨1의 대표 프레임에서 원하는 장면을 확인하고 원 비디오로 접근한다. 경우 3)은 경우 2)에서 요약 비디오 레벨1 대표 프레임에서 바로 원 비디오로 접근하기 어려운 경우, 더 자세한 정보를 얻기 위하여 가장 관심 있는 대표 프레임을 선택하고 그 대표 프레임 근처의 원 비디오 대표 프레임들을 대상으로 원하는 장면을 확인하고 원 비디오의 대표 프레임을 이용하여 원 비디오로 접근한다. 경우 4)와 경우 5)는 요약 비디오 레벨1의 재현에서 시작하고 경로는 위에서 설명한 경우와 유사하다.

<66> 본 발명의 기술 사상은 상기 바람직한 실시 예에 따라 구체적으로 기술되었으나, 상기한 실시예는 그 설명을 위한 것이며 그 제한을 위한 것이 아님을 주의하여야 한다. 또한, 본 발명의 기술 분야의 통상의 전문가라면 본 발명의 기술 사상의 범위 내에서 다양한 실시예가 가능함을 이해할 수 있을 것이다.

### 【발명의 효과】

<67> 이상에서 설명한 바와 같이 본 발명은,

<68>      요약 비디오를 통하여 비디오 전체 내용을 빠른 시간에 파악하고 요약 비디오의 각 비디오 구간의 대표 프레임 정보와 대표 음향 정보를 이용하여 효과적인 계층적 브라우징을 가능하게 한다. 또한 주문형 요약 비디오 기능을 통하여 내용에 따른 다양한 종류의 요약 비디오를 사용자에게 제공하는 효과가 있다.

**【특허청구범위】****【청구항 1】**

요약비디오를 기술하는 계층적 요약기술구조에 있어서,

적어도 하나 이상의 요약 비디오구간들의 요약비디오구간정보를 각각 기술하는 상기 요약 비디오구간 수만큼의 하이라이트 세그먼트기술구조를 포함하는 하이라이트레벨 기술구조와,

상기 하이라이트레벨의 의미를 기술하는 계층유형을 포함하고,

상기 하이라이트 세그먼트기술구조는 요약 비디오구간의 시간정보를 기술하거나 요약 비디오구간의 비디오 데이터정보를 기술하는 세그먼트 위치지정기술구조를 포함하여 이루어진 것을 특징으로 하는 계층적 요약기술구조.

**【청구항 2】**

제 1 항에 있어서, 상기 하이라이트레벨 기술구조는 다수의 하이라이트레벨을 포함하고, 상기 계층유형은 상기 하이라이트레벨의 레벨간의 관계를 더 포함한 것을 특징으로 하는 계층적 요약기술구조.

**【청구항 3】**

제 1 항 또는 제 2 항에 있어서, 상기 계층적 요약기술구조는,

요약비디오구간을 선택하는 데 기반이 된 사건들의 종류를 기술하는 요약기준기술  
구조를 더 포함한 것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 4】

제 1 항 또는 제 2 항에 있어서, 상기 하이라이트 세그먼트 기술구조는,  
비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조를 더 포함한  
것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 5】

제 4 항에 있어서, 상기 영상위치지정기술구조는, 대표프레임에 대한 시간구간정보  
를 기술하는 것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 6】

제 4 항에 있어서, 상기 영상위치지정기술구조는, 대표프레임의 영상데이터를 기술  
하는 것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 7】

제 1 항 또는 제 2 항에 있어서, 상기 하이라이트 세그먼트기술구조는, 비디오구간  
을 대표하는 음향 정보를 기술하는 대표음향기술구조를 더 포함한 것을 특징으로 하는

계층적 요약기술구조.

#### 【청구항 8】

제 7 항에 있어서, 상기 대표음향기술구조는, 대표음향에 대한 음향시간구간정보를 기술하는 것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 9】

제 7 항에 있어서, 상기 대표음향기술구조는, 대표음향의 음향데이터를 기술하는 것을 특징으로 하는 계층적 요약기술구조.

#### 【청구항 10】

제 1 항 또는 제 2 항에 있어서, 상기 하이라이트세그먼트기술구조는, 요약비디오 구간의 사건종류를 기술하는 세그먼트유형을 더 포함한 것을 특징으로 하는 계층적 요약 기술구조.

#### 【청구항 11】

제 1 항 또는 제 2 항에 있어서, 상기 하이라이트세그먼트기술구조는,  
비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조;  
비디오구간을 대표하는 음향 정보를 기술하는 대표음향기술구조; 및

요약비디오구간의 사건종류를 기술하는 세그먼트유형을 더 포함한 것을 특징으로 하는 계층적 요약기술구조.

### 【청구항 12】

적어도 하나 이상의 요약 비디오구간들의 요약비디오구간정보를 각각 기술하는 상기 요약 비디오구간 수만큼의 하이라이트 세그먼트기술구조를 포함하는 하이라이트레벨 기술구조와, 상기 하이라이트레벨의 의미를 기술하는 계층유형을 포함하고, 상기 하이라이트 세그먼트기술구조는, 요약 비디오구간의 시간정보를 기술하거나 요약 비디오구간의 비디오 데이터정보를 기술하는 세그먼트 위치지정기술구조와, 상기 요약비디오구간을 대표하는 프레임 정보를 기술하는 영상위치지정기술구조와, 상기 요약비디오구간을 대표하는 음향 정보를 기술하는 대표음향기술구조와, 상기 요약비디오구간의 사건종류를 기술하는 세그먼트유형구조를 포함하는 데이터가 기록된 컴퓨터로 읽을 수 있는 기록매체.

### 【청구항 13】

원 비디오를 입력받아 요약 기술구조에 따라서 요약 비디오기술 데이터를 생성하는 요약 비디오 기술데이터 생성방법에 있어서,  
원 비디오를 입력받고 분석하여 비디오 분석결과를 출력하는 비디오 분석단계와;  
요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의단계;  
상기 원 비디오 분석 결과와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내

용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택단계; 및

상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적 요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술단계를 포함하여 이루어진 것을 특징으로 하는 요약 비디오 기술데이터 생성방법.

#### 【청구항 14】

제 13 항에 있어서, 상기 비디오 분석단계는,  
원 비디오를 입력으로 받아들여 특징을 추출하고, 추출된 특징의 종류와 특징이  
검출되는 시간구간을 출력하는 특징추출단계와,  
상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주  
요 사건들을 검출하는 사건검출단계,  
상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여  
에피소드를 검출하는 에피소드 검출단계를 포함하여 이루어진 것을 특징으로 하는 요약  
비디오 기술 데이터 생성방법.

#### 【청구항 15】

제 13 항 또는 제 14 항에 있어서, 상기 요약규칙정의단계는 요약비디오구간을 선  
택하는데 기반이 되는 요약사건종류를 정의하여 상기 요약비디오기술단계에 제공하는 것

을 특징으로 하는 요약 비디오 기술 데이터 생성방법.

### 【청구항 16】

제 13 항 또는 제 14 항에 있어서, 상기 요약 비디오 구간정보를 입력받아 대표 프레임을 추출하여 상기 요약비디오기술단계에 제공하는 대표프레임 추출단계를 더 포함한 것을 특징으로 하는 요약 비디오 기술 데이터 생성방법.

### 【청구항 17】

제 16 항에 있어서, 상기 요약 비디오 구간정보를 입력받아 대표음향을 추출하여 상기 요약비디오기술단계에 제공하는 대표음향추출단계를 더 포함한 것을 특징으로 하는 요약 비디오 기술 데이터 생성방법.

### 【청구항 18】

원 비디오를 입력받아 요약 기술구조에 따라서 요약 비디오기술 데이터를 생성하는 요약 비디오 기술데이터 생성시스템에 있어서,

원 비디오를 입력받고 분석하여 비디오 분석결과를 출력하는 비디오 분석수단과;

요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의수단;

상기 원 비디오 분석 결과와 상기 요약 규칙을 입력받아 원 비디오로부터 비디오 내용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택수단; 및

상기 요약비디오 구간선택수단에서 정의된 요약비디오구간정보를 입력받아 계층적 요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술수단을 포함하여 이루어진 것을 특징으로 하는 요약 비디오 기술데이터 생성시스템.

#### 【청구항 19】

제 18 항에 있어서, 상기 비디오 분석수단은,  
원 비디오를 입력으로 받아들여 특징을 추출하고, 추출된 특징의 종류와 특징이  
검출되는 시간구간을 출력하는 특징추출수단과,  
상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주  
요 사건들을 검출하는 사건검출수단,  
상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여  
에피소드를 검출하는 에피소드 검출수단을 포함하여 이루어진 것을 특징으로 하는 요약  
비디오 기술 데이터 생성시스템.

#### 【청구항 20】

제 18 항 또는 제 19 항에 있어서, 상기 요약규칙정의수단은 요약비디오구간을 선  
택하는데 기반이 되는 요약사건종류를 정의하여 상기 요약비디오기술수단에 제공하는 것  
을 특징으로 하는 요약 비디오 기술 데이터 생성시스템.

**【청구항 21】**

제 18 항 또는 제 19 항에 있어서, 상기 요약비디오구간선택수단의 요약 비디오 구간정보를 입력받아 대표 프레임을 추출하여 상기 요약비디오기술수단에 제공하는 대표프레임 추출수단을 더 포함한 것을 특징으로 하는 요약 비디오 기술 데이터 생성시스템.

**【청구항 22】**

제 21 항에 있어서, 상기 요약비디오구간선택수단의 요약 비디오 구간정보를 입력 받아 대표음향을 추출하여 상기 요약비디오기술수단에 제공하는 대표음향추출수단을 더 포함한 것을 특징으로 하는 요약 비디오 기술 데이터 생성방법.

**【청구항 23】**

컴퓨터에,

입력되는 원 비디오에서 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출단계와,

상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출단계,

상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출단계

요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의단계;

상기 검출된 에피소드와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내용을

요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택단계; 및

상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적 요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술단계를 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

#### 【청구항 24】

컴퓨터에,

입력되는 원 비디오에서 특징을 추출하고, 추출된 특징의 종류와 특징이 검출되는 시간구간을 출력하는 특징추출수단과,

상기 특징의 종류와 특징이 검출되는 시간구간을 입력받아 원 비디오에 포함된 주요 사건들을 검출하는 사건검출수단,

상기 검출된 주요 사건들을 바탕으로 원 비디오를 이야기 흐름기반으로 구분하여 에피소드를 검출하는 에피소드 검출수단,

요약 비디오 구간을 선택하기 위한 요약 규칙을 정의하는 요약규칙 정의수단,

상기 검출된 에피소드와 상기 요약 규칙을 입력받아 원 비디오에서 비디오내용을 요약할 수 있는 비디오 구간을 선택하여 요약비디오구간정보를 구성하는 요약비디오 구간선택수단, 및

상기 요약비디오구간선택단계에서 정의된 요약비디오구간정보를 입력받아 계층적 요약기술구조를 가지는 비디오요약기술데이터를 생성하는 요약비디오기술수단을 기능시키

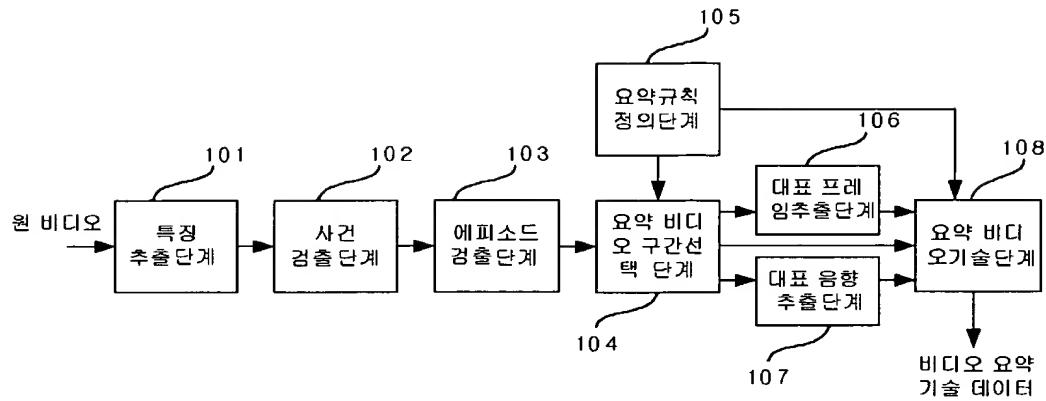
1019990043712

2000/9/2

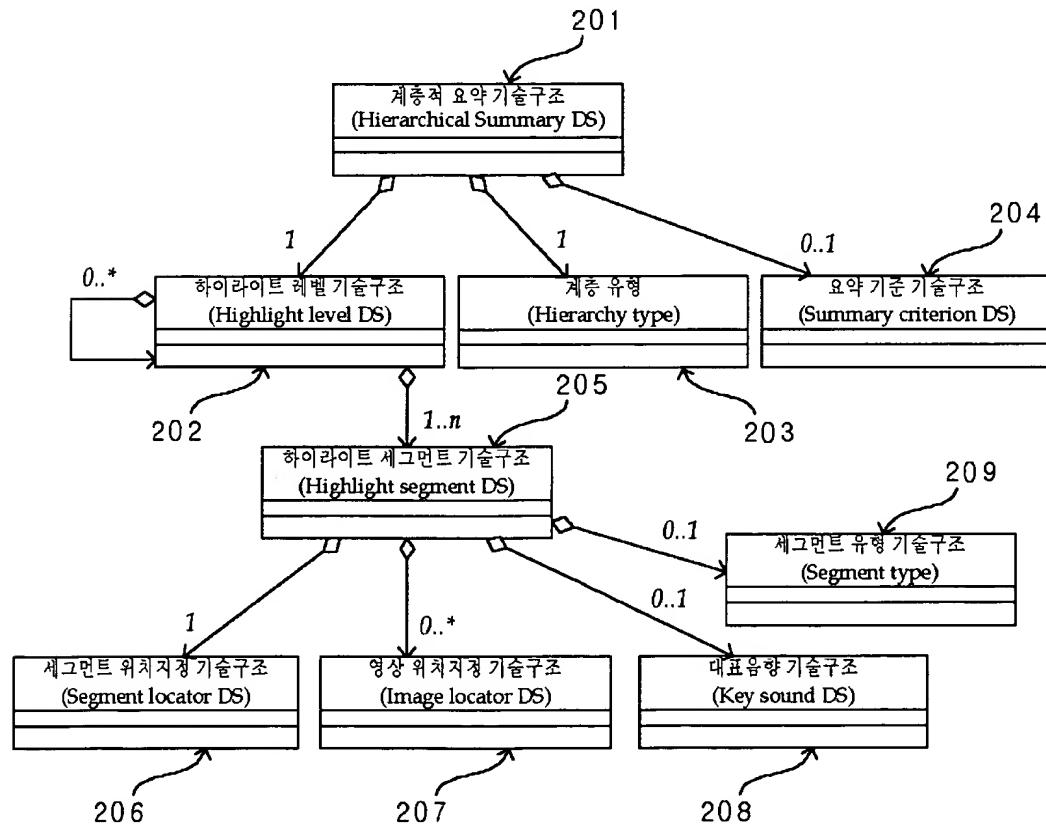
기 위한 컴퓨터로 읽을 수 있는 기록매체.

## 【도면】

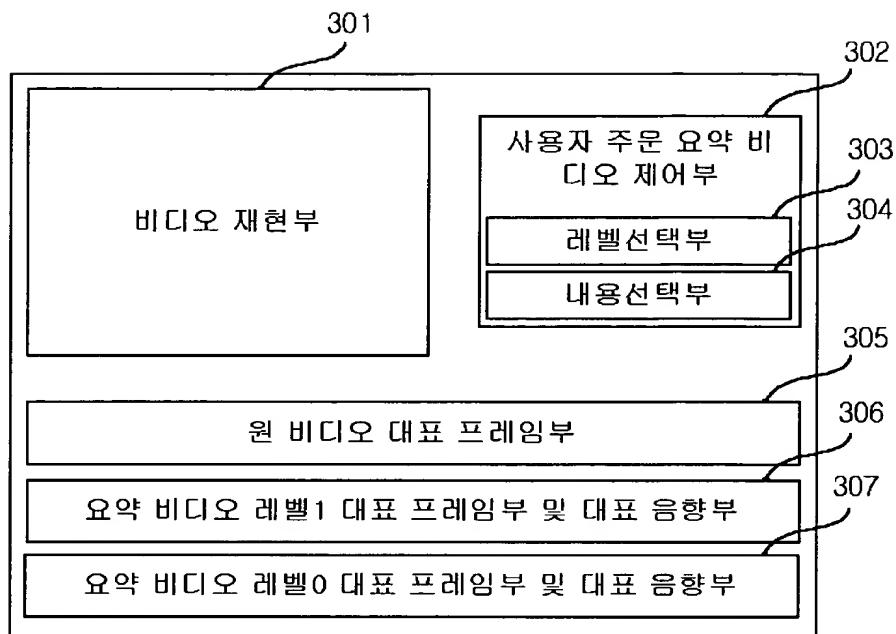
【도 1】



【도 2】



【도 3】



【도 4】

